Vectors and Matrices

Based on Lectures and "Intro to Linear Algebra"

$\theta\omega\theta$

Not in University of Cambridge
skipped some talks irrelevant to contents

E-mail: not telling you

Contents

PART

I

Complex Numbers

1 Definition

Definition 1.1. Construct $\mathbb{C}$ from $\mathbb{R}$ by adding i that $i^2 = -1$. Any $z \in \mathbb{C}$ is in the form

$$z = x + iy, x = \operatorname{Re} z, y = \operatorname{Im} z, x, y \in \mathbb{R}.$$

Addition and multiplication are defined by

$$z_1 + z_2 = (x_1 + x_2) + i(y_1 + y_2), z_1 z_2 = (x_1 + iy_1)(x_2 + iy_2).$$

The *conjugate* is defined by

$$\bar{z} = z* = x - iy.$$

The *modulus* is defined by

$$r = |z|, r \geqslant 0, r^2 = |z|^2 = z\bar{z} = x^2 + y^2.$$

The *argument* is defined by

$$z \neq 0 : \theta = \arg(z) \in \mathbb{R}, z = r(\cos\theta + i\sin\theta).$$

The values of θ in $(-\pi, \pi]$ are called the *principal values.*
Complex numbers can be plotted on an *Argand diagram.*

2 Basic Properties & Consequences

(1) $+, \times$ are commutative and associative,

$\mathbb{C}$ under $+$ is an abelian group,

$\mathbb{C}$ under $\times$ is an abelian group,

$\mathbb{C}$ is a field.

(2) **Fundamental Theorem of Algebra**: A polynomial with deg n with coefficients in $\mathbb{C}$ can be written as a product of n linear factors, has at least one solution in $\mathbb{C}$ and has n solutions connected with multiplicity.

(3) Parallelogram constructions.

(4)

$$|z_1||z_2| = |z_1 z_2|, |z_1 + z_2| \leqslant |z_1| + |z_2|.$$

Alternative forms:

$$|z_2 - z_1| \geqslant |z_2| - |z_1|, |z_2 - z_1| \geqslant ||z_2| - |z_1||.$$

(5) **De Moivre's Theorem**: $z^n = r^n(\cos n\theta + i\sin n\theta)$.

3 Exponential and Trigs in $\mathbb{C}$

Definition 3.1. Define $\exp, \cos, \sin$ on $\mathbb{C}$ by

$$\begin{aligned} \exp(z) &= e^z = \sum_{n=0}^{\infty} \frac{1}{n!} z^n, \\ \cos(z) &= \frac{1}{2}(e^{iz} + e^{-iz}) = 1 - \frac{1}{2!}z^2 + \frac{1}{4!}z^4 + \cdots, \\ \sin(z) &= \frac{1}{2i}(e^{iz} - e^{-iz}) = z - \frac{1}{3!}z^3 + \frac{1}{5!}z^5 + \cdots. \end{aligned}$$

These series converge for all $z \in \mathbb{C}$. Can be multiplied, rearranged, etc. Definitions reduce to familiar ones in the reals.

Proposition 3.1. $\forall z, w \in \mathbb{C}, e^z e^w = e^{z+w}; e^z e^{-z} = 1, (e^z)^n = e^{nz}, n \in \mathbb{Z}$.

Lemma 3.1. For $z = x + iy$:

(1) $e^z = e^x(\cos y + i \sin y)$.

(2) $\exp(z) \in \mathbb{C} \setminus \{0\}$.

(3) $e^z = 1 \Leftrightarrow z = 2\pi n i, n \in \mathbb{Z}$.

Definition 3.2 (Roots of unity). z is an Nth root of unity if $z^N = 1$.

We have

$$z^N = r^N e^{iN\theta} = 1 \iff r = 1, N\theta = 2n\pi \iff \theta = \frac{2n\pi}{N},$$

which gives N distinct solutions

$$z = \frac{2n\pi}{N} = \omega^n, \quad , n = 0, 1, \ldots, N-1.$$

ω^n lie one the vertices of a regular n-gon on the unit circle.

4 Logarithms and Complex powers

Definition 4.1. Define $w = \log z, z \in \mathbb{C} \wedge z \neq 0$ by $e^w = e^{\log z} = z$. Note that since exp is many-to-one, log is multi-valued.

$$\begin{aligned} z = re^{i\theta} = e^{\log r} e^{i\theta} = e^{\log r + i\theta} \\ \implies \boxed{\log z = \log r + i\theta = \log |z| + i \arg(z)} \end{aligned}$$

To make it single-valued, simply take the principal value.

Definition 4.2. Define *complex power* by

$$z^\alpha = e^{\alpha \log z}, \quad z, \alpha \in \mathbb{C}, z \neq 0.$$

Note that since $\arg z \to \arg z + 2n\pi \Rightarrow z^\alpha \to z^\alpha e^{2n\pi}$, it is generally multi-valued. This also reduces to common powers when $z, \alpha \in \mathbb{R}$.

Example.

$$i^i = e^{i\log i} = e^{i(0+i(\frac{\pi}{2}+2n\pi))} = e^{-(\frac{\pi}{2}+2n\pi)}.$$

5 Transformations, Lines, and Circles

- We have five elementary transformations:
  (1) $z \mapsto z + a$,
  (2) $z \mapsto \lambda z$,
  (3) $z \mapsto e^{i\alpha} z$,
  (4) $z \mapsto \bar{z}$,
  (5) $z \mapsto \frac{1}{z}$.
- General point of a line in $\mathbb{C}$ through z_0 and parallel to w:

$$z = z + \lambda w, \lambda \in \mathbb{R} \text{ or } \bar{w}z - w\bar{z} = \bar{w}z_0 - w\bar{z}_0.$$

- General point of a circle in $\mathbb{C}$ with centre c and radius ρ:

$$z = c + \rho e^{i\theta} \text{ or } |z - c| = \rho \text{ or } |z|^2 - \bar{c}z - c\bar{z} = \rho^2 - |c|^2.$$

- Stereographic projection.

PART II

Vectors in 3 Dimensions

6 Vector addition and scalar multiplication

Definition 6.1 (scalar multiplication)**.** Given $\mathbf{a}$, and scalar $\lambda \in \mathbb{R}$, define $\lambda\mathbf{a}$ to be the position vector of A' on the line OA with length $|\lambda\mathbf{a}| = |\lambda|\,|\mathbf{a}|$. Direction depends on the sign of λ.

Define span $\{\mathbf{a}\} = \{\lambda\mathbf{a} : \lambda \in \mathbb{R}\}$. If $\mathbf{a} \neq 0$, then span$\{\mathbf{a}\}$ is the entire line through O and A.

Define $\mathbf{a} \parallel \mathbf{b}$ if and only if either $\mathbf{a} = \lambda\mathbf{b}$ or $\mathbf{b} = \lambda\mathbf{a}$. Allow $\lambda = 0$, so $\forall \mathbf{a}, \mathbf{0} \parallel \mathbf{a}$. Also allow $\lambda < 0$.

Definition 6.2 (vector addition)**.** Give $\mathbf{a}, \mathbf{b}$, if $\mathbf{a} \nparallel \mathbf{b}$, construct a parallelogram $OACB$ and define $\mathbf{c} = \mathbf{a} + \mathbf{b}$.

If $a \parallel b$, then $\mathbf{a} = \alpha\mathbf{u}, \mathbf{b} = \beta\mathbf{u}$, where $\mathbf{u}$ is a unit vector and $\mathbf{a} + \mathbf{b} = (\alpha + \beta)\mathbf{u}$.

Given $\mathbf{a}, \mathbf{b}, \ldots, \mathbf{c}$, we have a linear combination

$$\alpha\mathbf{a} + \beta\mathbf{b} + \cdots + \gamma\mathbf{c}$$

for any $\alpha, \beta, \ldots, \gamma \in \mathbb{R}$.

Define span $\{\mathbf{a}, \mathbf{b}, \ldots, \mathbf{c}\} = \{\alpha\mathbf{a} + \beta\mathbf{b} + \cdots + \gamma\mathbf{c} : \alpha, \beta, \ldots, \gamma \in \mathbb{R}\}$. In 3d case, if $\mathbf{a} \nparallel \mathbf{b}$, then span$\{\mathbf{a}, \mathbf{b}\}$ is a plane through O, A, B.

Here are some properties:

- $\forall \mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{a} + \mathbf{0} = \mathbf{0} + \mathbf{a} = \mathbf{a}$, this says that $\mathbf{0}$ is the identity for addition.
- $\exists - \mathbf{a}, \mathbf{a} + (-\mathbf{a}) = (-\mathbf{a}) + \mathbf{a} = \mathbf{0}$. This says $-\mathbf{a}$ is the inverse of $\mathbf{a}$ under addition.
- $\forall \mathbf{a}, \mathbf{b}, \mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$, this says that vector addition is commutative.
- $\forall \mathbf{a}, \mathbf{b}, \mathbf{c}, (\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$, this says that vector addition is associative.

Hence, the set of vectors with addition form an abelian group.

Relation with scalars:

- $\lambda(\mathbf{a} + \mathbf{b}) = \lambda\mathbf{a} + \lambda\mathbf{b}$.
- $(\lambda + \mu)\mathbf{a} = \lambda\mathbf{a} + \mu\mathbf{a}$.
- $(\lambda\mu)\mathbf{a} = \lambda(\mu\mathbf{a})$.

7 Dot product

Definition 7.1 (dot product)**.** Give $\mathbf{a}, \mathbf{b}$, let θ be the angle between them, define $\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos\theta$. Note that θ is defined unless $\mathbf{a} = \mathbf{0}$, in which case we define $\mathbf{a} \cdot \mathbf{b} = 0$. $\mathbf{a} \perp \mathbf{b} \Leftrightarrow \mathbf{a} \cdot \mathbf{b} = 0 \Leftrightarrow \theta = \frac{\pi}{2} \bmod \pi$ when θ is defined. Allow $\mathbf{a}$ or $\mathbf{b} = 0$, so $\mathbf{a} \parallel \mathbf{0} \wedge \mathbf{a} \perp \mathbf{0}$.

For $\mathbf{a} \neq \mathbf{0}$, $|\mathbf{b}| \cos\theta$ is the component of $\mathbf{b}$ along $\mathbf{a}$.

$$|\mathbf{b}| \cos\theta = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}|} = \mathbf{u} \cdot \mathbf{b}.$$

By resolving $\mathbf{b}$ along and perpendicular to $\mathbf{a}$, we get

$$\mathbf{b} = \mathbf{b}_{\parallel} + \mathbf{b}_{\perp}.$$

Properties:

- $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}, \mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|^2 \geqslant 0, = 0$ iff $\mathbf{a} = \mathbf{0}$.
- $(\lambda\mathbf{a}) \cdot \mathbf{b} = \lambda(\mathbf{a} \cdot \mathbf{b}) = \mathbf{a} \cdot (\lambda\mathbf{b})$.
- $\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}$.

8 Vector cross product

Definition 8.1. Given $\mathbf{a}, \mathbf{b}$, let θ be the angle between them, wrt a unit vector $\mathbf{n}$ normal to the plane they span. Define $\mathbf{a} \wedge \mathbf{b}$ or $\mathbf{a} \times \mathbf{b}$ as $|\mathbf{a}||\mathbf{b}| \sin\theta\mathbf{n}$. $\mathbf{0}$ case is similar.

This is the *vector area* of the parallelogram generated by $\mathbf{a}, \mathbf{b}$. Note that $\mathbf{a} \wedge \mathbf{b} = \mathbf{a} \wedge \mathbf{b}_{\perp}$.

Properties:

- $\mathbf{a} \wedge \mathbf{b} = \mathbf{b} \wedge \mathbf{a}$.
- $(\lambda\mathbf{a}) \wedge \mathbf{b} = \lambda(\mathbf{a} \wedge \mathbf{b}) = \mathbf{a} \wedge (\lambda\mathbf{b})$.
- $\mathbf{a} \wedge (\mathbf{b} + \mathbf{c}) = \mathbf{a} \wedge \mathbf{b} + \mathbf{a} \wedge \mathbf{c}$.
- $\mathbf{a} \wedge \mathbf{b} = \mathbf{0}$ if and only if $\mathbf{a} \parallel \mathbf{b}$.
- $\mathbf{a} \wedge \mathbf{b} \perp \mathbf{a} \wedge \perp \mathbf{b}$.

9 Orthonormal Bases and Components

Choose $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ that are *orthonormal*. That is, they are of unit lengths and $\mathbf{e}_i \cdot \mathbf{e}_j = 0, i \neq j \in \{1, 2, 3\}$, which is equivalent to choose cartesian axes along the directions. Then $\{\mathbf{e}_i\}$ is a basis and $\forall \mathbf{a} \in \mathbb{R}^3$,

Lecture 4.

$$\mathbf{a} = \sum_{i=1}^{3} a_i \mathbf{e}_i.$$

Each component a_i is uniquely determined by $a_i = \mathbf{e}_i \cdot \mathbf{a}$.

By this spirite, we can write

$$\mathbf{a} = (a_1, a_2, a_3) = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}.$$

Scalar product in this form can be written as

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = a_1 b_1 + a_2 b_2 + a_3 b_3, \quad |\mathbf{a}| = a_1^2 + a_2^2 + a_3^2..$$

For vector products, choose this basis that it is also *right-handed*:

$$\mathbf{e}_i \times \mathbf{e}_{i+1} = \mathbf{e}_{i+2}.$$

Then

$$\begin{aligned} \mathbf{a} \times \mathbf{b} &= (a_1 \mathbf{e_1} + a_2 \mathbf{e}_2 + a_3 \mathbf{e}_3)(b_1 \mathbf{e_1} + b_2 \mathbf{e}_2 + b_3 \mathbf{e}_3) \\ &= (a_3 b_2 - a_2 b_3)\mathbf{e}_1 + (a_3 b_1 - a_1 b_3)\mathbf{e}_2 + (a_1 b_2 - a_2 b_1)\mathbf{e}_3. \end{aligned}$$

10 Triple products

10.1 Scalar triple product

Definition 10.1. Define scalar triple product by

$$[\mathbf{a}, \mathbf{b}, \mathbf{c}] = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}).$$

This is the volumn of the parallelepiped with bases $\mathbf{b}, \mathbf{c}$ and side $\mathbf{a}$.

Remark. $\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})$ is a "signed" volumn. If $\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) > 0$ then $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ is called a *right-handed set*. $\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})$ if and only if $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are coplanar, e.g., $\mathbf{c} = \alpha \mathbf{a} + \beta \mathbf{b} \in \operatorname{span}\{\mathbf{a}, \mathbf{b}\}$.

In components,

$$\begin{aligned} \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = a_1 b_2 c_3 - a_1 b_3 c_2 & \\ + a_2 b_3 c_1 - a_2 b_1 c_3 &= \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}. \\ + a_3 b_1 c_2 - a_3 b_2 c_1 & \end{aligned}$$

10.2 Vector triple product

Definition 10.2. Define the vector triple product by $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$. Note that $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ does not necessarily give the same result as $(\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$.

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}. \tag{10.1}$$

We have the following identities:

Proposition 10.1.

$$\begin{aligned}
&\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) + \mathbf{b} \times (\mathbf{c} \times \mathbf{a}) + \mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = \mathbf{0} \\
&(\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{d}) = [\mathbf{a}, \mathbf{b}, \mathbf{d}]\mathbf{c} - [\mathbf{a}, \mathbf{b}, \mathbf{c}]\mathbf{d} \\
&(\mathbf{a} \times \mathbf{b}) \cdot ((\mathbf{c} \times \mathbf{d}) \times (\mathbf{e} \times \mathbf{f})) = [\mathbf{a}, \mathbf{b}, \mathbf{d}][\mathbf{c}, \mathbf{e}, \mathbf{f}] - [\mathbf{a}, \mathbf{b}, \mathbf{c}][\mathbf{d}, \mathbf{e}, \mathbf{f}] \\
&(\mathbf{b} \times \mathbf{c}) \cdot (\mathbf{a} \times \mathbf{d}) + (\mathbf{c} \times \mathbf{a}) \cdot (\mathbf{b} \times \mathbf{d}) + (\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = 0
\end{aligned}$$

11 Lines, Planes, and Vector equations

Vectors are defined as position vectors from O. But the definition of addition enables us to use them to describe displacements between points.

11.1 Lines

General point on a line through $\mathbf{a}$ through $\mathbf{u}$:

$$\begin{aligned}
&\mathbf{r} = \mathbf{a} + \lambda\mathbf{u}, \qquad \lambda \in \mathbb{R} \quad \text{The parametric form.} \\
&\mathbf{u} \times \mathbf{r} = \mathbf{u} \times \mathbf{a}, \qquad\qquad \text{Cross form.}
\end{aligned}$$

Proposition 11.1. Any vector equation of the form $\mathbf{u} \times \mathbf{r} = \mathbf{c}$ represents a line.

Proof. $\mathbf{u} \times \mathbf{r} = \mathbf{c} \Rightarrow \mathbf{u} \cdot (\mathbf{u} \times \mathbf{r}) = \mathbf{u} \cdot \mathbf{c} \Leftrightarrow \mathbf{u} \cdot \mathbf{c} = 0$. If $\mathbf{u} \cdot \mathbf{c} \neq 0$ then the equation is inconsistent. If $\mathbf{u} \cdot \mathbf{c} = 0$, then note that

$$\mathbf{u} \times (\mathbf{u} \times \mathbf{c}) = (\mathbf{u} \cdot \mathbf{c})\mathbf{u} - (\mathbf{u} \cdot \mathbf{u})\mathbf{c} = -|\mathbf{u}|^2\mathbf{c}.$$

Hence $\mathbf{a} = -(\mathbf{u} \times \mathbf{c})/|\mathbf{u}|^2$ is a solution, and thus it represents a line. □

11.2 Planes

General point on a plane through $\mathbf{a}$ with directions $\mathbf{u}, \mathbf{v}$ in the plane($\mathbf{u} \nparallel \mathbf{v}$):

$$\begin{aligned}
&\mathbf{r} = \mathbf{a} + \lambda\mathbf{u} + \mu\mathbf{v}, \qquad \lambda, \mu \in \mathbb{R} \quad \text{Parametric form,} \\
&\mathbf{n} \cdot \mathbf{r} = k = \mathbf{n} \cdot \mathbf{a}, \qquad \mathbf{n} = \mathbf{u} \times \mathbf{v} \quad \text{Dot form.}
\end{aligned}$$

The component of $\mathbf{r}$ along $\mathbf{n}$ is

$$\frac{\mathbf{n} \cdot \mathbf{r}}{|\mathbf{n}|} = \frac{k}{|\mathbf{n}|}.$$

11.3 Other vector equations

(1) $|\mathbf{r}|^2 + \mathbf{r} \cdot \mathbf{a} = k \Leftrightarrow \left|\mathbf{r} + \frac{1}{2}\mathbf{a}\right|^2 = k + \frac{1}{4}|\mathbf{a}|^2$, a sphere with centre $-\frac{1}{2}\mathbf{a}$ and radius $\sqrt{k + \frac{1}{4}|\mathbf{a}|^2}$, provided $k > -\frac{1}{4}|\mathbf{a}|^2$.

(2) $\mathbf{r} + \mathbf{a} \times (\mathbf{b} \times \mathbf{r}) = \mathbf{c} \Leftrightarrow \mathbf{r} + (\mathbf{a} \cdot \mathbf{r})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{r} = \mathbf{c}$. Dot with $\mathbf{a}$:

$$\mathbf{a} \cdot \mathbf{r} = \mathbf{a} \cdot \mathbf{c} \Longrightarrow (1 - \mathbf{a} \cdot \mathbf{b})\mathbf{r} = \mathbf{c} - (\mathbf{a} \cdot \mathbf{c})\mathbf{b}.$$

If $\mathbf{a} \cdot \mathbf{b} \neq 1$, then there is a unique solution

$$\mathbf{r} = \frac{1}{1 - \mathbf{a} \cdot \mathbf{b}}(\mathbf{c} - (\mathbf{a} \cdot \mathbf{c})\mathbf{b}),$$

which is a point.

If $\mathbf{a} \cdot \mathbf{b} = 1$ and RHS$\neq 0$, then it is inconsistent.

If $\mathbf{a} \cdot \mathbf{b}$ and $\mathbf{c} - (\mathbf{a} \cdot \mathbf{c})\mathbf{b} = \mathbf{0}$, then

$$(\mathbf{a} \cdot \mathbf{r} - \mathbf{a} \cdot \mathbf{c})\mathbf{b} = \mathbf{0}.$$

Hence it is a plane.

12 Index notation and the summation convention

12.1 Components, $\delta \& \epsilon$

Write vectors $\mathbf{a}, \mathbf{b}, \ldots$ in terms of components $a_i, b_i, \ldots$ wrt an orthonormal right-handed basis $\{\mathbf{e}_i\}$. Indices $i, j, \ldots$ take values $1, 2, 3$.

For example, if $\mathbf{c} = \alpha\mathbf{a} + \beta\mathbf{b}$, then $c_i = [\alpha\mathbf{a} + \beta\mathbf{b}]_i = \alpha a_i + \beta b_i$, for $i = 1, 2, 3$. i is called a *free index*.

Hence

- $\mathbf{a} \cdot \mathbf{b} = \displaystyle\sum_{i=1}^{3} a_i b_i$.
- $\mathbf{x} = \mathbf{a} + (\mathbf{b} \cdot \mathbf{c})\mathbf{d} \Leftrightarrow x_j = a_j + \left(\displaystyle\sum_{k=1}^{3} b_k c_k\right) d_j$.

Definition 12.1 (Kronecker delta).

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

We see that $\delta_{ij} = \delta_{ji}$ and also

$$\begin{pmatrix} \delta_{11} & \delta_{12} & \delta_{13} \\ \delta_{21} & \delta_{22} & \delta_{23} \\ \delta_{31} & \delta_{32} & \delta_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Using this, $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$

Definition 12.2 (Levi-Civita epsilon).

$$\epsilon_{ijk} = \begin{cases} 1 & \text{if } (i, j, k) \text{ is an even permutation of } (1, 2, 3), \\ -1 & \text{if } (i, j, k) \text{ is an odd permutation of } (1, 2, 3), \\ 0 & \text{else.} \end{cases}$$

We have $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1, \epsilon_{321} = \epsilon_{213} = \epsilon_{132} = -1$. ϵ_{ijk} is totally anti-symmetric: exchanging any pair of indices produces a change in sign.

Then

$$\mathbf{e}_i \times \mathbf{e}_j = \sum_k \epsilon_{ijk} \mathbf{e}_k$$

and

$$\begin{aligned}\mathbf{a}\times\mathbf{b} &= \left(\sum_i a_i\mathbf{e}_i\right)\times\left(\sum_j b_j\mathbf{e}_j\right)\\ &= \sum_{ij} a_ib_j\mathbf{e}_i\times\mathbf{e}_j\\ &= \sum_{ij} a_ib_j\sum_k \epsilon_{ijk}\mathbf{e}_k = \sum_{ijk} a_ib_j\epsilon_{ijk}\mathbf{e}_k\end{aligned}$$

so

$$\boxed{(\mathbf{a}\times\mathbf{b})_k = \sum_{ij}\epsilon_{ijk}a_ib_j}.$$

12.2 Summation convention

With components and index notation, indices that appear twice in a given term are usually summed over. In the summation convention, we omit the sum signs for repeated indices. i.e., the sum is understood.

Example. (i) In $a_i\delta_{ij} = a_1\delta_{1j} + a_2\delta_{2j} + a_3\delta_{3j} = a_j$, since Σ_i is understood.

(ii) $\mathbf{a}\cdot\mathbf{b} = \delta_{ij}a_ib_j = a_ib_i$. Σ_{ij}, Σ_i are understood.

(iii) $(\mathbf{a}\times\mathbf{b})_i = \epsilon_{ijk}a_jb_k$, Σ_{jk} is understood.

(iv) $[\mathbf{a},\mathbf{b},\mathbf{c}] = \epsilon_{ijk}a_ib_jc_k$, Σ_{ijk} is understood.

(v) $\delta_{ii} = \delta_{11} + \delta_{22} + \delta_{33} = 3$.

(vi)

$$\begin{aligned}[(\mathbf{a}\cdot\mathbf{c})\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\mathbf{c}]_i &= (\mathbf{a}\cdot\mathbf{c})b_i - (\mathbf{a}\cdot\mathbf{b})c_i\\ &= a_jc_jb_i - a_jb_jc_i.\end{aligned}$$

Σ_j is understood.

Here are the rules of summation convention.

(1) An index occuring exactly once in any given term must appear once in every term in an equation, and it can take any value in $1, 2, 3$, a *free* index.

(2) An index occuring exactly twice in a given term is summed over. A *repeated*, *contracted*, or *dummy* index.

(3) No index can occur more than twice in any given term.

12.3 Applications

We can use this to prove the vector triple product identity.

Proof. Write the huge sum in summation convention:

$$\begin{aligned}[\mathbf{a}\times(\mathbf{b}\times\mathbf{c})]_i &= \epsilon_{ijk}a_j(\mathbf{b}\times\mathbf{c})_k\\ &= \epsilon_{ijk}a_j\epsilon_{kpq}b_pc_q\\ &= (\epsilon_{ijk}\epsilon_{kpq})a_jb_pc_q.\end{aligned}$$

Notice that

$$\epsilon_{ijk}\epsilon_{kpq} = \delta_{ip}\delta_{jq} - \delta_{iq}\delta_{jp} \qquad (*)$$

see next subsection. So

$$[\mathbf{a}\times(\mathbf{b}\times\mathbf{c})]_i = \delta_{ip}\delta_{jq}a_jb_pc_q - \delta_{iq}\delta_{jp}a_jb_pc_q.$$

Notice also that $a_i\delta_{ij} = a_j$, so

$$[\mathbf{a}\times(\mathbf{b}\times\mathbf{c})]_i = a_qb_ic_q - a_jb_jc_i = (\mathbf{a}\cdot\mathbf{c})b_i - (\mathbf{a}\cdot\mathbf{b})c_i.$$

Hence the equation 10.1 is proved. □

12.4 $\epsilon\,\epsilon$ identity

Proposition 12.1. $\epsilon_{ijk}\epsilon_{pqk} = \delta_{ip}\delta_{jq} - \delta_{iq}\delta_{jp} = \epsilon_{kij}\epsilon_{kpq}$.

Proof. Notice that LHS and RHS are both anti-symmetric, so both vanish when i, j or p, q take the same value. Inspection shows that[1] it suffices to show the cases $i = p = 1 \wedge j = q = 2$ or $i = q = 1, j = p = 2$ and all other index changings that give non-zero results. □

[1] *Think carefully here.*

Proposition 12.2. $\epsilon_{ijk}\epsilon_{pjk} = 2\delta_{ip}$.

Proof. Take $q = j$ in the above equation:

$$\epsilon_{ijk}\epsilon_{pjk} = \delta_{ip}\delta_{jj} - \delta_{ij}\delta_{jp} = 3\delta_{ip} - \delta_{ip} = 2\delta_{ip}.$$

□

Proposition 12.3. $\epsilon_{ijk}\epsilon_{ijk} = k$.

Proposition 12.4.

$$\begin{aligned}\epsilon_{ijk}\epsilon_{pqr} = {} & \delta_{ip}\delta_{jq}\delta_{kr} - \delta_{jp}\delta_{iq}\delta_{kr} \\ & + \delta_{jp}\delta_{kq}\delta_{ir} - \delta_{kp}\delta_{jq}\delta_{ir} \\ & + \delta_{kp}\delta_{iq}\delta_{jr} - \delta_{ip}\delta_{kq}\delta_{jr}.\end{aligned}$$

Proof. Total anti-symmetry[2] in i, j, k and independently in p, q, r implies LHR, RHS agree up to an overall factor. To check the factor is 1, consider $i = p = 1, j = q = 2, k = r = 3$. □

[2] *This simplifies most of the process and leaves only one case to check.*

PART III

Vectors in General

13 Vectors in $\mathbb{R}^n$

13.1 Definition and basic properties

Lecture 6

Definition 13.1. Regard vectors as sets of components, and let

$$\mathbb{R}^n = \{\mathbf{x} = (x_1, \ldots, x_n) : x_i \in \mathbb{R}\}.$$

Define:

- Addition: $\mathbf{x} + \mathbf{y} = (x_1 + y_1, \ldots, x_n + y_n)$,

- Scalar multiplication: $\lambda\mathbf{x} = (\lambda x_1, \ldots, \lambda x_n)$.
- Linear combinations: $\lambda\mathbf{x} + \mu\mathbf{y}$,
- Parallel: $\mathbf{x} \parallel \mathbf{y} \Leftrightarrow \mathbf{x} = \lambda\mathbf{y} \vee \mathbf{y} = \lambda\mathbf{x}$.
- Inner Product(Scalar product): $\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} x_i y_i$.

Properites of inner product:

(1). Symmetric: $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$.

(2). Bilinear:

$$\begin{aligned}(\lambda\mathbf{x} + \lambda'\mathbf{x}') \cdot \mathbf{y} &= \lambda\mathbf{x} \cdot \mathbf{y} + \lambda'\mathbf{x}' \cdot \mathbf{y},\\ \mathbf{x} \cdot (\mu\mathbf{y} + \mu'\mathbf{y}') &= \mu\mathbf{x} \cdot \mathbf{y} + \mu'\mathbf{x}' \cdot \mathbf{y}.\end{aligned}$$

(3). Positive definite: $\mathbf{x} \cdot \mathbf{x} \geqslant 0$, with $=$ holds if and only if $\mathbf{x} = \mathbf{0}$.

13.2 Norm of a vector

Definition 13.2. The *norm* of a vector $\mathbf{x}$ is denoted as $|\mathbf{x}|$ with $|\mathbf{x}|^2 = \mathbf{x} \cdot \mathbf{x}$.
$\mathbf{x}, \mathbf{y}$ are called *orthogonal* if $\mathbf{x} \cdot \mathbf{y} = 0$, denote as $\mathbf{x} \perp \mathbf{y}$.

The *standard basis* of $\mathbb{R}^n$ is

$$e_i = (0, \ldots, 1, \ldots, 0)$$

with 1 on the ith position. So that

$$\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i$$

and $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$. i.e., standard basis is orthogonal.

13.3 Cauchy-Schwarz and Triangle inequalities

Proposition 13.1 (Cauchy-Schwarz)**.**

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \quad |\mathbf{x} \cdot \mathbf{y}| \leqslant |\mathbf{x}||\mathbf{y}|$$

with equality if and only if $\mathbf{x} \parallel \mathbf{y}$.

General deductions:

(i). Setting $\mathbf{x} \cdot \mathbf{y} = |\mathbf{x}||\mathbf{y}| \cos\theta$, we can define angle θ between $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

(ii). We have the *triangle inequality*:

$$|\mathbf{x} + \mathbf{y}| \leqslant |\mathbf{x}| + |\mathbf{y}|.$$

Proof. If $\mathbf{y} = \mathbf{0}$, then the result is immediate. If not, consider

$$\begin{aligned}|\mathbf{x} - \lambda\mathbf{y}| &= (\mathbf{x} - \lambda\mathbf{y})(\mathbf{x} - \lambda\mathbf{y})\\ &= |\mathbf{x}|^2 - 2\lambda\mathbf{x} \cdot \mathbf{y} + \lambda^2|\mathbf{y}|^2 \geqslant 0.\end{aligned}$$

This is a real equation of λ with at most one root, so

$$(-2\mathbf{x}\cdot\mathbf{y})^2 - 4|\mathbf{x}|^2|\mathbf{y}|^2 \leqslant 0 \iff |\mathbf{x}\cdot\mathbf{y}| \leqslant |\mathbf{x}||\mathbf{y}|.$$

Equality holds if and only if $\mathbf{x} = \lambda\mathbf{y}$.
Note also that for triangle inequality:

$$\begin{aligned}|\mathbf{x}+\mathbf{y}|^2 &= |\mathbf{x}|^2 + 2\mathbf{x}\cdot\mathbf{y} + |\mathbf{y}|^2\\ &\leqslant |\mathbf{x}|^2 + 2|\mathbf{x}||\mathbf{y}| + |\mathbf{y}|^2\\ &= (|\mathbf{x}|+|\mathbf{y}|)^2,\end{aligned}$$

as required. $\square$

13.4 Inner Products and Cross products

Inner product in $\mathbb{R}^n$ can be written as

$$\mathbf{a}\cdot\mathbf{b} = \delta_{ij}a_ib_j, \quad \text{by summation convention.}$$

For $n = 3$, it matches geometrical definition.

We can also define cross product in component definition. In 3d we have

$$(\mathbf{a}\times\mathbf{b})_i = \epsilon_{ijk}a_jb_k,$$

and in n dimensions we have $\epsilon_{ij\dots l}$ which is totally anti-symmetric. But there are only two a_ib_j so we cannot use this to define vector product in general.

However, in $\mathbb{R}^2$ we have ϵ_{ij} with $\epsilon_{12} = -\epsilon_{21} = 1$, so can use this to define a new scalar product

$$[\mathbf{a},\mathbf{b}] = \epsilon_{ij}a_ib_j = a_1b_2 - a_2b_1.$$

Geometrically, this the (signed) area of parallelogram formed by $\mathbf{a},\mathbf{b}$ and

$$|[\mathbf{a},\mathbf{b}]| = |\mathbf{a}||\mathbf{b}|\sin\theta.$$

Compare with $[\mathbf{a},\mathbf{b},\mathbf{c}] = \epsilon_{ijk}a_ib_jc_k$.

14 Vector Spaces

14.1 Axioms, span, and subspaces

Definition 14.1. Let V be a set of objects called *vectors* with operation

$$\begin{aligned}\mathbf{v}+\mathbf{w} \in V &\quad \forall \mathbf{v},\mathbf{w}\in V\\ \lambda\mathbf{v}\in V &\quad \forall \mathbf{v}\in V, \lambda\in\mathbb{R}.\end{aligned}$$

Then V is called a *real vector space* if

(i). V with $+$ is an abelian group.

(ii). $\lambda(\mathbf{v}+\mathbf{w}) = \lambda\mathbf{v}+\lambda\mathbf{w}$

(iii). $(\lambda+\mu)\mathbf{v} = \lambda\mathbf{v}+\mu\mathbf{v}$

(iv). $\lambda(\mu\mathbf{v}) = (\lambda\mu)\mathbf{v}$

(v). $1\mathbf{v} = \mathbf{v}$.

Example. Let $V = \{f : [0,1] \to \mathbb{R} : f \wedge f(0) = f(1) = 0\}$. By smooth we mean f is differentiable infinitely many times. Then V is a real vector space with $+$ defined as $(f+g)(x) = f(x) + g(x)$ and $(\lambda f)(x) = \lambda(f(x))$. Then all axioms apply.

Definition 14.2. A *subspace* of a real vector space V is a subset $U \subseteq V$ that is also a vector space.

Lecture 7

Remark. A non-empty subset is a subspace if and only if $\forall v, w \in U, \lambda v + \mu u \in U$.

For any vectors $v_1, \ldots, v_r \in V$, their *span* $\operatorname{span}\{v_1, \ldots, v_r\} = \{\lambda_1 v_1 + \cdots + \lambda_r v_r : v_i \in \mathbb{R}\}$ is a subspace. V and $\{0\}$ are subspaces of V.

Example. A line or plane through O is a subspace in $\mathbb{R}^3$, but a line or plane that does not contain $\mathbf{0}$ is not a subspace.

14.2 Linear dependence and independence

For $v_1, \ldots, v_r \in V$, a real vector space, consider a linear relation

$$\lambda_1 v_1 + \cdots + \lambda_r v_r = 0. \tag{*}$$

If $(*) \Rightarrow \lambda_i = 0$, then the vectors form a *linearly independent set*. They obey only the trivial linear relation.

If $(*)$ holds with at least $\lambda_k \neq 0$, then the vectors form a *linearly dependent* set. They obey a non-trivial linear relation.

Example. In $\mathbb{R}^2$, $\{(1,0),(0,1),(0,2)\}$ is linearly dependent.
We cannot express $(1,0)$ in terms of the others.

Several facts:

- Any set containing 0 is linearly dependent.
- In $\mathbb{R}^3$, $\{\mathbf{a}\}$ is linearly independent if and only if $\mathbf{a} \neq \mathbf{0}$.
- $\{\mathbf{a}, \mathbf{bc}\}$ is linearly independent if $[\mathbf{a}, \mathbf{b}, \mathbf{c}] \neq 0$. Since if

  $$\alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c} = 0,$$

  then dotting with $\mathbf{b} \times \mathbf{c}$ we get $\alpha[\mathbf{a}, \mathbf{b}, \mathbf{c}] = 0 \Rightarrow \alpha = 0$. Similarly $\beta = 0, \gamma = 0$.

14.3 Inner products

This is an additional structure on a real vector space V, that can also be characterised by axioms or key properties.

For $v, w \in V$, denote inner product by

$$v \cdot w \text{ or } (v, w) \in \mathbb{R}.$$

Require this satisfies 1. it's symmetric, 2. it's bilinear, 3. it is positive definite.

Definition of length or norm and deductions such as Cauchy-Schwarz inequality depend just on these properties.

Example. Consider space of functions

$$V = \{f : [0,1] \to \mathbb{R} : f \text{ smooth} \wedge f(0) = f(1) = 0\}.$$

Define an inner product by

$$(f, g) = \int_0^1 f(x)g(x)\,\mathrm{d}x.$$

which has properties 123. Cauchy-Schwarz holds:

$$|(f, g)| \leqslant \|f\|\,\|g\|$$

with $\|f\|^2 = (f, f)$. i.e.

$$\left|\int_0^1 f(x)g(x)\,\mathrm{d}x\right| \leqslant \left(\int_0^1 f(x)\,\mathrm{d}x\right)^{1/2} \left(\int_0^1 g(x)\,\mathrm{d}x\right)^{1/2}.$$

Lemma 14.1. In any real vector space V with an inner product, if $v_1, v_2, \ldots, v_r$ are non-zero and orthogonal vectors, then they are linearly independent.

Proof. If

$$\sum_i \alpha_i v_i = 0,$$

then

$$(v_j, \sum_i \alpha_i v_i) = 0 \iff \alpha_j = 0.$$

□

15 Bases and dimension

Definition 15.1. For a vector space V, a *basis* is a set

$$\mathfrak{B} = \{e_1, \ldots, e_n\}$$

such that

(i) $\mathfrak{B}$ spans V. i.e., $\forall v \in V$,

$$v = \sum_{i=1}^n v_i e_i.$$

(ii) $\mathfrak{B}$ is linearly independent.

Given (ii), the coefficients v_i in (i) are unique, since

$$\sum_i v_i e_i = \sum_i v_i' e_i \iff v_i - v_i' = 0 \iff v_i = v_i'.$$

Example. Standard basis for $\mathbb{R}^n$ consists of

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Many other bases can be chosen.

Theorem 15.1. If $\{e_1, \dots, e_n\}$, $\{f_1, \dots, f_m\}$ are bases for a real vector space V, then $m = n$.

Proof. We have
$$f_a = \sum_i A_{ai} e_i,$$
$$e_i = \sum_a B_{ia} f_a$$
for $A_{ai}, B_{ia} \in \mathbb{R}$. Hence
$$f_a = \sum_i A_{ai} \sum_b B_{ib} f_b$$
$$= \sum_b \sum_i A_{ai} B_{ib} f_b$$
But the coefficients are unique, so
$$\sum_i A_{ai} B_{ib} = \delta_{ab}.$$
Similarly,
$$\sum_a B_{ia} A_{aj} = \delta_{ij}.$$
Now,
$$\sum_{i,a} A_{ai} B_{ia} = \sum_a \delta_{aa} = m = \sum_i \delta_{ii} = n, .$$
□

Definition 15.2. The number of vectors in any basis is the *dimension* of the vector space.

Remark. $\{0\}$ is called the *trivial* vector space and has dimension 0. Lecture 8
The steps in the proof of basis theorem are within scope of this course, but the proof without prompts non-examinable. The same applies to the following:

Proposition 15.1. Let V be a vector space with finite subsets $Y = \{w_1, \dots, w_m\}$ and $X = \{u_1, \dots, u_k\}$, with Y spans V and X linearly independent. Then
$$k \leqslant \dim V \leqslant m.$$
And

(1) A basis can be found as a subset of Y be discarding vectors as necessary.

(2) X can be extended to a basis by adding vectors from Y as necessary.

Proof. (1) If Y is linearly independent, then Y is a basis, and $m = n = \dim V$. If Y is not, then
$$\sum_{i=1}^{m} \lambda_i w_i = 0,$$

where λ_i are not all zero. wlog, can take $\lambda_m \neq 0$. Then

$$w_m = -\frac{1}{\lambda_m}\sum_{i=1}^{m-1}\lambda_i w_i,$$

so $\operatorname{span} Y = \operatorname{span}(Y \setminus \{w_m\}) = \operatorname{span} Y'$. We repeat, until a basis is obtained.
(2) If X spans V then X is a basis $k = n = \dim V$. If not $\exists u_{k+1}$ not in $\operatorname{span} X$. Consider

$$\sum_{i=1}^{k+1}\mu_i u_i = 0$$

and thus $\mu_i = 0, \forall i \in \{1, 2, \dots, k+1\}$. Hence

$$X' = X \cup \{u_{k+1}\}$$

is linearly independent.
Furthermore, we can choose $u_{k+1} \in Y$ since otherwise $\operatorname{span} X = V$, #. Repeat this until a basis is achieved. The process stops since Y is finite. □

In this course, we will deal only with finite-dimensional spaces, except examples mentioned.

Example. $V = \{f : [0,1] \to \mathbb{R} : f \text{ smooth } \wedge f(0) = f(1) = 0\}$. Note that

$$s_n(x) = \sqrt{2}\sin(n\pi x)$$

belong to V and

$$(s_n, s_m) = 2\int_0^1 \sin(n\pi x)\sin(m\pi x)\,\mathrm{d}x = \delta_{nm},$$

so these functions are orthonormal and thus linearly independent. So V is infinite-dimensional.

16 Vectors in $\mathbb{C}^n$

16.1 Introduction and definitions

Let $\mathbb{C}^n = \{\mathbf{z} = (z_1, \dots, z_n) : z_j \in \mathbb{C}\}$ and define

- Addition: $\mathbf{z} + \mathbf{w} = (z_1 + w_1, \dots, z_n + w_n)$,
- Scalar multiplication: $\lambda\mathbf{z} = (\lambda z_1, \dots, \lambda z_n)$.

If scalars $\lambda, \mu \in \mathbb{R}$, then $\mathbb{C}^n$ is a real vector space, and axioms apply.

If $\lambda, \mu \in \mathbb{C}$, $\mathbb{C}^n$ is a complex vector space. The same axioms hold, and definitions of linear combinations, linear dependence/independence, bases, dimension are generalised to $\mathbb{C}$.

The distinction between real and complex scalars is important.

Example. $\mathbf{z} = (z_1, \dots, z_n) \in \mathbb{C}^n$ with $z_j = x_j + iy_j, x_j, y_j \in \mathbb{R}$. Then

$$\mathbf{z} = \sum_j x_j\mathbf{e}_j + \sum_j y_j\mathbf{f}_j,$$

a linear combination of $\mathbf{e}$, the usual standard basis in $\mathbb{R}^n$, and $\mathbf{f}_j = (0, \dots, i, \dots, 0)$.

We can see that $\{\mathbf{e}_1, \ldots, \mathbf{e}_n, \mathbf{f}_1, \ldots, \mathbf{f}_n\}$ is a basis for $\mathbb{C}^n$ as a *real* vector space, so it has dimension $2n$.

However,

$$\mathbf{z} = \sum_j z_j \mathbf{e}_j$$

is a *complex* linearl combination, so the basis of $\mathbb{C}^n$ as a *complex* vector space is simply $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ and the dimension is n over $\mathbb{C}$.

From now on we will view $\mathbb{C}^n$ as a complex vector space unless mentioned otherwise.

16.2 Inner product

The inner product on $\mathbb{C}^n$ is defined by

$$(\mathbf{z}, \mathbf{w}) = \sum_{j=1}^{n} \bar{z}_j w_j.$$

It has the followin properties:

(1) It is *hermitian*: $(\mathbf{w}, \mathbf{z}) = \overline{(\mathbf{z}, \mathbf{w})}$.

(2) It is linear/anti-linear: $(\mathbf{z}, \lambda\mathbf{w} + \lambda'\mathbf{w}') = \lambda(\mathbf{z}, \mathbf{w}) + \lambda'(\mathbf{z}, \mathbf{w}')$. But $(\lambda\mathbf{z} + \lambda'\mathbf{z}', w) = \bar{\lambda}(\mathbf{z}, \mathbf{w}) + \bar{\lambda}'(\mathbf{z}, \mathbf{w}')$.

(3) Positive definite: $(\mathbf{z}, \mathbf{z}) \in \mathbb{R} \wedge \geqslant 0$. $= 0$ if and only if $\mathbf{z} = \mathbf{0}$.

Define the norm of $\mathbf{z}$ to be

$$|\mathbf{z}| \geqslant 0, |\mathbf{z}|^2 = (\mathbf{z}, \mathbf{z}).$$

Also define $\mathbf{z}, \mathbf{w} \in \mathbb{C}^n$ to be *orthogonal* if $(\mathbf{z}, \mathbf{w}) = 0$.

Note that the standard basis for $\mathbb{C}^n$ is orthonormal. That is,

$$(\mathbf{e}_j, \mathbf{e}_k) = \delta_{jk}.$$

Example. Complex inner product of $\mathbb{C}^1$ is

$$(z, w) = \bar{z}w.$$

Let $z = a_1 + ia_2, w = b_1 + ib_2$, and considers $\mathbf{a} = (a_1, a_2), \mathbf{b} = (b_1, b_2) \in \mathbb{R}^2$. Then

$$\bar{z}w = a_1b_1 + a_2b_2 + i(a_1b_2 - a_2b_1) = \mathbf{a} \cdot \mathbf{b} + i[\mathbf{a}, \mathbf{b}],$$

recover 2 scalar products in $\mathbb{R}^2$.

PART

IV

Matrices and Linear Maps

17 Introduction

17.1 Definitions

Lecture 9

Definition 17.1. A *linear map* or *linear transformation* is a function $T : V \to W$ between V with $\dim V = n$ and W with $\dim W = m$ such that

$$T(\lambda x + \mu y) = \lambda T(x) + \mu T(y)$$

where $x, y \in V$ and $\lambda, \mu \in \mathbb{R}$ or $\mathbb{C}$, depending on whether V, W are complex vector spaces.
$x' = T(x) \in W$ is called the *image* of x under T. Define

$$\begin{aligned} \operatorname{Im}(T) &= \{x' \in W : \exists x \in V, x' = T(x)\}, \\ \ker(T) &= \{x \in V : T(x) = 0 \in W\}. \end{aligned}$$

Remark. A linear map is determined by its action on a basis. We have

$$T\left(\sum_i x_i e_i\right) = \sum_i x_i T(e_i).$$

Lemma 17.1. $\operatorname{Im}(T), \ker(T)$ are subspace of W, V respectively.

Proof. Note that $0 \in \operatorname{Im}(T)$ and $\forall x', y' \in \operatorname{Im}(T)$, let $T(x) = x', T(y) = y'$, then $\lambda x' + \mu y' = \lambda T(x) + \mu T(y) = T(\lambda x + \mu y) \in \operatorname{Im}(T)$, so it is a subspace. $\ker(T)$ is proved similarly. □

Example. (1) Zero linear map: $T : V \to W$ that $T(v) = 0$. We have $\operatorname{Im}(T) = \{0\}$ and $\ker(T) = V$.

(2) Identity map: $T : V \to V$ that $T(v) = v$. $\operatorname{Im}(T) = V, \ker(T) = \{0\}$.

(3) Let $V = W = \mathbb{R}^3$ and $T(x) = x'$ where

$$\begin{aligned} x_1' &= 3x_1 + x_2 + 5x_3, \\ x_2^x &= -x_1 - 2x_3, \\ x_3' &= 2x_1 + x_2 + 3x_3. \end{aligned}$$

Then T is indeed a linear map and

$$\operatorname{Im}(T) = \left\{\lambda \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}\right\}, \quad \ker(T) = \left\{\lambda \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix}\right\}.$$

Note that $\operatorname{Im}(T)$ is a plane and $\ker(T)$ is a line.

17.2 Rank and Nullity

Define $\operatorname{rank}(T) = \dim \operatorname{Im}(T)$ and $\operatorname{null}(T) = \dim \ker(T)$.

Theorem 17.2. For $T : V \to W$ a linear map, then

$$\operatorname{rank}(T) + \operatorname{null}(T) = \dim V.$$

Proof. Let $\{e_1, \ldots, e_k\}$ be a basis of $\ker(T)$. Extend this by $e_{k+1}, \ldots, e_n$ to a basis of V. Claim that $\mathcal{B} = \{T(e_{k+1}), \ldots, T(e_n)\}$ is a basis of $\operatorname{Im}(T)$. The result clearly follows.
Indeed, $\mathcal{B}$ spans $\operatorname{Im}(T)$ since $\forall x \in V, x = \sum_i x_i e_i$ and

$$T(x) = \sum_{i=1}^{n} x_i T(e_i) = \sum_{i=k+1}^{n} x_i T(e_i).$$

Suppose

$$\sum_{i=k+1}^{n} \lambda T(e_i) = 0.$$

Thus

$$\begin{aligned} T\left(\sum_{i=k+1}^{n} x_i e_i\right) &= 0 \\ \implies \sum_{i=k+1}^{n} x_i e_i &\in \ker(T) \\ \implies \sum_{i=k+1}^{n} x_i e_i &= \sum_{i=1}^{k} x_i e_i \\ \impliedby -\sum_{i=1}^{k} x_i e_i + \sum_{i=k+1}^{n} x_i e_i &= 0 \end{aligned}$$

$$\implies x_i = 0 \text{ for } i = 1, \ldots, n.$$

Hence $\mathcal{B}$ is linearly independent and thus it is a base. □

18 Geometrical Examples

18.1 Rotations

In $\mathbb{R}^2$, rotations about $\mathbf{0}$ through θ is defined by

$$\begin{aligned} \mathbf{e}_1 &\mapsto \mathbf{e}_1' = \cos(\theta)\mathbf{e}_1 + \sin(\theta)\mathbf{e}_2, \\ \mathbf{e}_2 &\mapsto \mathbf{e}_2' = -\sin(\theta)\mathbf{e}_1 + \cos(\theta)\mathbf{e}_2. \end{aligned}$$

In $\mathbb{R}^3$, can extend so that $\mathbf{e}_3 \mapsto \mathbf{e}_3$

To generalise to arbitrary rotations of θ along axis $\mathbf{n}$, where $\mathbf{n}$ is a unit vector, resolve horizontally and vertically:

$$\mathbf{x} = \mathbf{x}_\parallel + \mathbf{x}_\perp,$$

where $\mathbf{x}_\parallel = (\mathbf{x} \cdot \mathbf{n})\mathbf{n}$. Then,

$$\begin{aligned} \mathbf{x}_\parallel &\mapsto \mathbf{x}_\parallel, \\ \mathbf{x}_\perp &\mapsto \cos(\theta)\mathbf{x}_\perp + \sin(\theta)\mathbf{n} \times \mathbf{x}, \end{aligned}$$

by considering the plane perpendicular to $\mathbf{n}$. Note that $|\mathbf{x}_\perp| = |\mathbf{x} \times \mathbf{n}|$, so the result follows by comparing with $\mathbb{R}^2$ by regarding $\mathbf{e}_1$ as $\mathbf{x}_\perp$ and $\mathbf{e}_2$ as $\mathbf{n} \times \mathbf{x}$.

Hence,

$$\begin{aligned}\mathbf{x} &\mapsto (\mathbf{n}\cdot\mathbf{x})\mathbf{x} + \cos\theta(\mathbf{x} - (\mathbf{n}\cdot\mathbf{x})\mathbf{x}) + \sin\theta\mathbf{n}\times\mathbf{x} \\ &= \boxed{\cos\theta\mathbf{x} + (1-\cos\theta)(\mathbf{n}\cdot\mathbf{x})\mathbf{n} + \sin\theta\mathbf{n}\times\mathbf{x}.}\end{aligned}$$

18.2 Reflections and Projections

For a plane with unit normal vector $\mathbf{n}$, define projectin of $\mathbf{x}$ on the plane as

$$\begin{aligned}\mathbf{x}_\parallel &\mapsto \mathbf{0}, \\ \mathbf{x}_\perp &\mapsto \mathbf{x}_\perp.\end{aligned}$$

i.e.,

$$\mathbf{x} \mapsto \mathbf{x}' = \mathbf{x} - (\mathbf{n}\cdot\mathbf{x})\mathbf{n}.$$

Define reflection of $\mathbf{x}$ wrt the plane by

$$\begin{aligned}\mathbf{x}_\parallel &\mapsto -\mathbf{x}_\parallel, \\ \mathbf{x}_\perp &\mapsto \mathbf{x}_\perp.\end{aligned}$$

i.e.,

$$\mathbf{x} \mapsto \mathbf{x}' = \mathbf{x} - 2(\mathbf{n}\cdot\mathbf{x})\mathbf{n}.$$

Same applies to $\mathbb{R}^2$ by replacing plane by line.

18.3 Dilations

Given scale factors $\alpha, \beta, \gamma > 0$. Define a *dilation* along axes by

$$\begin{aligned}\mathbf{e}_1 &\mapsto \mathbf{e}_1' = \alpha\mathbf{e}_1, \\ \mathbf{e}_2 &\mapsto \mathbf{e}_2' = \beta\mathbf{e}_2, \\ \mathbf{e}_3 &\mapsto \mathbf{e}_3' = \gamma\mathbf{e}_3.\end{aligned}$$

Then $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 \mapsto \mathbf{x}' = \alpha x_1\mathbf{e}_1 + \beta x_2\mathbf{e}_2 + \gamma x_3\mathbf{e}_3$.

18.4 Shears

Let $\mathbf{a}, \mathbf{b}$ be orthogonal unit vectors in $\mathbb{R}^3$, and λ a real parameter. Define a *shear*

$$\mathbf{x} \mapsto \mathbf{x}' = \mathbf{x} + \lambda\mathbf{a}(\mathbf{x}\cdot\mathbf{b}).$$

Notice that $\mathbf{a} \mapsto \mathbf{a}$ and $\mathbf{b} \mapsto \mathbf{b} + \lambda\mathbf{a}$. Definition holds the same way in $\mathbb{R}^2$.

19 Matrices as linear maps

19.1 Definitions

Lecture 10

Consider a linear map $T : \mathbb{R}^n \to \mathbb{R}^m$, with bases $\{\mathbf{e}_i\}_{i=1}^n, \{\mathbf{f}_a\}_{a=1}^m$, of the form

$$T(\mathbf{x}) = \mathbf{x}', \quad \mathbf{x} = \sum_{i=1}^n x_i\mathbf{e}_i, \mathbf{x}' = \sum_{a=1}^m x_a' f_a.$$

Linearity of T implies we can specify T using $T(\mathbf{e}_i) = \mathbf{e}'_i = \mathbf{C}_i \in \mathbb{R}^m$. Take these $\mathbf{C}_i$ as *columns* of an $m \times n$ array or *matrix* M with rows $\mathbf{R}_a \in \mathbb{R}^n$.

$$M = \begin{pmatrix} \uparrow & & \uparrow \\ \mathbf{C}_1 & \cdots & \mathbf{C}_n \\ \downarrow & & \downarrow \end{pmatrix} = \begin{pmatrix} \leftarrow & \mathbf{R}_1 & \rightarrow \\ & \cdots & \\ \leftarrow & \mathbf{R}_m & \rightarrow \end{pmatrix}.$$

M has *entries* $M_{ai} \in \mathbb{R}$ where a labels rows and i labels columns. Thus we have

$$M_{ai} = (\mathbf{C}_i)_a = (\mathbf{R}_a)_i.$$

The action of T is then given by

$$\mathbf{x}' = M\mathbf{x},$$

defined by $x'_a = \sum_{i=1}^n M_{ai}x_i = M_{ai}x_i$, by summation convention. In column vector:

$$\begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_m \end{pmatrix} = \begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1n} \\ M_{21} & M_{22} & \cdots & M_{2n} \\ \vdots & \vdots & & \vdots \\ M_{m1} & M_{m2} & \cdots & M_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} M_{1i}x_i \\ M_{2i}x_i \\ \vdots \\ M_{mi}x_i \end{pmatrix}.$$

Indeed, this matrix does represent T, for

$$\begin{aligned} \mathbf{x}' &= T(x_i\mathbf{e}_i) = x_iT(\mathbf{e}_i) = x_i\mathbf{C}_i \\ \Longrightarrow x'_a &= x_i(\mathbf{C}_i)_a = M_{ai}x_i. \end{aligned}$$

Now we can regard properties of T as properties of M. For example,

$$\begin{aligned} &\operatorname{Im}(T) = \operatorname{Im}(M) = \operatorname{span}\{\mathbf{C}_1, \ldots, \mathbf{C}_n\}; \\ &x'_a = M_{ai}x_i = (\mathbf{R}_a)_ix_i = \mathbf{R}_a \cdot \mathbf{x}; \\ &\ker T = \ker M = \{\mathbf{x} : \mathbf{R}_a \cdot \mathbf{x} = 0 \text{ for all } a.\}. \end{aligned}$$

19.2 Examples

Example. (1) The *zero map* $\mathbb{R}^n \to \mathbb{R}^m$ corresponds to the *zero matrix*.

(2) Identity map corresponds to I where $I_{ij} = \delta_{ij}$, called the *unit matrix*.

(3) Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ with $\mathbf{x}' = T(\mathbf{x}) = M\mathbf{x}$ where

$$M = \begin{pmatrix} 3 & 1 & 5 \\ -1 & 0 & -2 \\ 2 & 1 & 3 \end{pmatrix}.$$

We get

$$\begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{pmatrix} 3x_1 + x_2 + 5x_3 \\ -x_1 - 2x_3 \\ 2x_1 + x_2 + 3x_3 \end{pmatrix}.$$

Hence if we let

$$\mathbf{C}_1 = \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix}, \mathbf{C}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 5 \\ -2 \\ 3 \end{pmatrix},$$

we get

$$\operatorname{Im} T = \operatorname{Im} M = \operatorname{span}\{\mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3\} = \operatorname{span}\{\mathbf{C}_1, \mathbf{C}_2\}.$$

Here we have

$$\mathbf{R}_1 = \begin{pmatrix} 3 & 1 & 5 \end{pmatrix}, \mathbf{R}_2 = \begin{pmatrix} -1 & 0 & 2 \end{pmatrix}, \mathbf{R}_3 = \begin{pmatrix} 2 & 1 & 3 \end{pmatrix},$$

hence $\mathbf{R}_2 \times \mathbf{R}_3 = \begin{pmatrix} 2 & -1 & -1 \end{pmatrix} = \mathbf{u}$, where infact $\mathbf{u} \perp \mathbf{R}_1$. Hence

$$\ker T = \ker M = \{\lambda \mathbf{u} : \lambda \in \mathbb{R}\}.$$

(4) Now we turn to study rotations in $\mathbb{R}^2$ and $\mathbb{R}^3$. The matrix wrt rotation of angle θ in $\mathbb{R}^2$ is

$$R = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

For $\mathbb{R}^3$, note that

$$\begin{aligned} \mathbf{x}' &= \cos\theta \mathbf{x} + (1-\cos\theta)(\mathbf{n}\cdot\mathbf{x})\mathbf{n} + \sin\theta \mathbf{n}\times\mathbf{x} \\ \Longrightarrow x_i' &= \cos\theta x_i + (1-\cos\theta) n_j x_j x_i - \sin\theta \epsilon_{ijk} x_j n_k = R_{ij} \\ \Longrightarrow R_{ij} &= \cos\theta \delta_{ij} + (1-\cos\theta) n_i n_j - \sin\theta \epsilon_{ijk} n_k. \end{aligned}$$

(5) Dilations. We have

$$M = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix}.$$

(6) Reflections. To find the matrix H wrt the reflection in plane with normal vector $\mathbf{n}$, consider

$$\begin{aligned} x_i' &= x_i - 2(x_j n_j) n_i = H_{ij} x_j \\ \Longrightarrow H_{ij} &= \delta_{ij} - 2 n_i n_j. \end{aligned}$$

(7) Shear. We have

$$\begin{aligned} \mathbf{x}' &= \mathbf{x} + \lambda(\mathbf{b}\cdot\mathbf{x})\mathbf{a} \\ \Longrightarrow x_i' &= x_i + \lambda(b_j x_j) a_i = S_{ij} x_j \\ \Longrightarrow S_{ij} &= \delta_{ij} + \lambda a_i b_j. \end{aligned}$$

19.3 Matrix of a General Linear Map $V \to W$

Definition 19.1. Consider $T : V \to W$, between general (real or complex) vector spaces, with dim n, m respectively. Choose $\{e_i\}_{i=1}^n, \{f_a\}_{a=1}^m$ as bases of V, W. Define the matrix of T wrt these bases is defined as an $m \times n$ array with entries $M_{ai} \in \mathbb{R}$ or $\mathbb{C}$ defined by

$$T(e_i) = \sum_{a=1}^m f_a M_{ai}.$$

Then $x' = T(x) \Leftrightarrow x_a' = M_{ai} x_i$.

Entries in column i of M are components of $T(e_i)$ wrt basis f_a.

Remark. Given choices of bases $\{e_i\}, \{f_a\}$, V is identified with $\mathbb{R}^n$ and W is identified with $\mathbb{R}^m$, and T is identified with an $m \times n$ matrix M.